Commercial Disclosure

Dr. Hillman has a financial interest in the APM based on a contractual agreement between Sensimetrics, Inc. (R&D for the initial version of the APM) and KayPENTAX, Inc. (manufacturer of the APM).

Ambulatory Monitoring of Daily Voice Use

Robert E. Hillman

Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital
Boston, MA
Communication Sciences and Disorders, MGH Institute of Health Professions, and Surgery and Health Sciences and Technology, Harvard Medical School

Cambridge, MA

Daryush D. Mehta

Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital
Boston, MA

Abstract

Many common voice disorders are chronic or recurring conditions likely to result from faulty and/or abusive patterns of vocal behavior. Such behaviorally based disorders can be difficult to assess accurately in the clinical setting and potentially could be much better characterized by long-term ambulatory monitoring of vocal function as individuals engage in their typical daily activities. Ambulatory monitoring also could provide new insight into the actual role of voice use in common disorders and missing quantitative data on what constitutes normal levels of daily voice use for different groups of individuals, activities, and occupations. This report describes the motivation, previous development efforts, current state-of-the-art technology, and future directions in ambulatory monitoring of voice use.

There is growing interest in developing medical devices for monitoring the function of body systems while individuals engage in their normal daily activities (vocation, avocation, etc.) to help facilitate the detection/diagnosis of abnormalities that may not be readily apparent during routine clinical examinations. Versions of systems exist for ambulatory monitoring of heart, respiratory, gastrointestinal, and brain function. One major impetus for expanding ambulatory monitoring capabilities is the desire to provide remote surveillance of health conditions/status and facilitate prevention and/or timely intervention, particularly as an aging population strives to live independently.

One area that has received recent attention has been in the development of devices for ambulatory monitoring of daily voice use. This area of focus has been motivated by the fact that many common voice disorders are chronic or recurring conditions likely to result from faulty and/or abusive patterns of vocal behavior, which produce phonotraumatic vocal fold lesions (e.g., vocal fold nodules) and/or result in functional dysphonia and vocal fatigue (Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1989). A higher risk for developing these types of disorders has been linked to occupations associated with high levels of voice use such as singers, counselors, social workers, teachers (primary or secondary school), lawyers, clergy, telemarketers, sales people, and healthcare workers (Verdolini & Ramig, 2001). Because voice use is considered to play a major role in the etiology of many common voice disorders, clinicians focus a great deal of attention on attempting to evaluate and modify how patients typically use their voices. Such efforts currently are limited by a reliance on patient self-reporting and self-monitoring. It is obvious that these approaches are very subjective and likely to be unreliable, particularly given that patterns of voice use (and misuse) become highly habituated and somewhat automatic and therefore are carried out below an individual's level of consciousness. There is actually a paucity of objective information about the actual role of daily voice use in the etiology of voice disorders, as well as minimal quantitative information about what constitutes normal levels of daily voice use (e.g., levels of voice use in different occupations; Cheyne, Hanson, Genereux, Stevens, & Hillman, 2003).

There are a variety of voice-use-related behaviors that are believed to contribute to voice disorders, particularly those associated with phonotrauma, including excessive loudness, inappropriate pitch, and perhaps phonating (speaking) for excessive lengths of time (Verdolini, Rosen, & Branski, 2005). Treatment of these disorders, which involves the identification and modification of such behaviors, would be greatly enhanced by the ability to unobtrusively monitor and quantify such behaviors through the measurement of distributions of intensity (sound pressure level; SPL), fundamental frequency (f0), and phonation duration over the course of a patient's typical day. Such information would enable clinicians to precisely pinpoint the location and duration of abusive and/or inappropriate behaviors and to specifically target these behaviors for modification. With appropriate enhancements, the monitoring device also could have great potential for improving treatment of voice-use-related disorders. Specifically, by providing real-time ambulatory biofeedback to patients (using an earphone or vibrating piece), abusive vocal behaviors could be signaled and then, hopefully, reduced.

Voice use is believed to be an important (but poorly understood) factor in how patients recover from laryngeal surgery. It is difficult for many patients to completely cease talking following laryngeal surgery, and some clinicians have been led to instruct patients to use their voices in ways that are intended to reduce the forces of phonation (reduced intensity, reduced talking time, etc.) and thus promote healing. A device that would allow ongoing monitoring of voice use during the critical post-operative period could improve patient compliance with these instructions and better ensure successful surgical recovery while offering a possible new mechanism for understanding how voice use affects the healing process.

Daily voice use is likely to vary widely from individual to individual, depending upon factors such as age, profession, living environment, and perhaps gender and ethnic group. A device that is capable of unobtrusively monitoring vocal function will enable the acquisition of important new information about the vocal demands (and/or vocal endurance) that are typically required for various types of normal speakers to carry out their daily activities (e.g., different professions). Such objective normative information is critical to developing a more clinically useful understanding of what constitutes "safe limits" for voice use and, conversely, identifying vocal functioning that places the voice user at increased risk for the development of voice-use-related vocal pathology. Clinicians could use the data obtained by such a device to better counsel patients with voice disorders about whether or not they would be able to meet the vocal demands of their job, or if they should plan to modify or adjust their job-related activities to conserve vocal function.

It has been long acknowledged that clinical approaches to the diagnosis and treatment of voice disorders could be greatly improved by the availability of instrumentation that provides continuous unobtrusive monitoring of important vocal parameters for an extended period of time, i.e., across the span of a typical day. There have been several attempts starting in the 1970s to develop such devices (Airo, Olkinuora, & Sala, 2000; Buekers, Bierens, Kingma, & Marres, 1995; Masuda, Ikeda, Manako, & Komiyam, 1993; Ohlsson, Brink, & Lofqvist, 1989; Ryu, Komiyama, Kannae, & Watanabe, 1983; Watanabe, Komiyama, Ryu, & Kannae, 1984; Watanabe, Shin, Oda, Fukaura, & Komiyama, 1987), which were generally referred to as "voice accumulators" (Buekers et al., 1995; Ohlsson et al.) or "speech accumulators" (Ryu et al.). While these earlier devices served to demonstrate the general proof of concept for ambulatory monitoring of voice use, as well as the potential benefits of such a device to clinical and research endeavors, they all appeared to have characteristics that would have significantly limited their use as clinical devices (e.g., obtrusive physical design, overly coarse sampling of vocal parameters, limited number of vocal parameters). It also is important to note that none of these devices has been made commercially available.

Some of the earlier devices (Ohlsson et al., 1989; Watanabe et al., 1987) did not record voice intensity and thus were not suited for application to voice disorders where excessive intensity is thought to be an important etiological factor. The study by Buekers and colleagues (1995) is probably the most relevant in that it provides both a proof of concept for this approach and some quantitative data that profiles both phonation duration and intensity for different professions. Repeated measures for 2 subjects showed day-to-day consistency in both the total duration of phonation and the intensity distribution of that phonation. This is important to establish for individual subjects and across subject groups if voice accumulators are to be used to track changes as a function of different interventions such as biofeedback or speech therapy protocols. Clear differences also were demonstrated as a function of occupation, which is consistent with observations that individuals in high-voice-use jobs (e.g., kindergarten teachers) have more frequent voice complaints than those in low-voice-use jobs (e.g., bookkeepers). Another interesting observation confirmed in other studies is that the actual total phonation time tends to be much less than speakers are likely to 'self-report' (Ohlsson et al.; Watanabe et al., 1987). In the Ohlsson and colleagues study, subjects estimated their voice use constituted 62%-71% of their time, while in fact their actual phonation time accounted for only about 7% of their time.

State-of-the-Art Technology

Efforts to improve ambulatory voice monitoring technology for research and clinical purposes have continued over the past decade. Our group (Cheyne, 2002; Cheyne et al., 2003; Hillman, Heaton, Masaki, Zeitels, & Cheyne, 2006; Zañartu et al., 2009) and others (Švec, Popolo, & Titze, 2003; Švec, Titze, & Popolo, 2005) have shown that a miniature accelerometer (ACC; Model BU-7135, Knowles Corp.) currently offers the best potential as a phonation sensor for long-term monitoring of vocal function because it is small and can be worn unobtrusively at the base of the neck (just above the sternal notch), has sensitivity and dynamic range characteristics that make it well suited to sensing vocal fold vibration, is relatively immune to other environmental sounds, has low power consumption, and produces a voice-related signal that is not filtered by the vocal tract, which, compared to a microphone signal, makes it easier to process, is more robust for sensing disordered phonation, and alleviates confidentiality concerns (speech audio is not recorded).

Two ACC-based ambulatory monitoring systems were developed in the early part of the past decade. One, by Titze and colleagues (Popolo, Švec, & Titze, 2005), utilized a personal digital assistant (PDA) as the data acquisition platform to create a proprietary research device. Unfortunately, design decisions that were specific to this PDA make implementation on newer platforms challenging. Data obtained from teachers with this device have produced valuable new information about voice use (Nix, Švec, Laukkanen, & Titze, 2007; Titze, Hunter, & Švec, 2007) and provided a test bed for developing theoretical concepts about vocal dose (sometimes referred to as vocal loading; Rantala, Vilkman, & Blogu, 2002; Švec et al., 2003; Titze, Švec, &

Popolo, 2003). The other ACC-based system was the result of a prolonged development effort by our group that culminated in the Ambulatory Phonation Monitor (APM), which is the first commercially-available device for research and clinical use (Model 3200, KayPENTAX, Lincoln Park, NJ). Figure 1 illustrates typical placement of the ACC for the APM device at the anterior neck region above the sternal notch. Both of these ACC-based devices unobtrusively collect long-term data on f0, SPL, and phonation duration while the wearer engages in typical daily activity.

Figure 1. Illustration of the commercially-available APM Model 3200 worn by a speaker. The inset shows the size of the accelerometer (mounted on a soft silicone pad) relative to a US penny.

Results from ambulatory voice monitoring for f0 and SPL can be expressed as average values and ranges, and/or displayed as distributions (histograms) across selected time periods (often an entire day). Phonation duration typically is reported as a percentage of the total selected monitoring period spent voicing (percent phonation time) and also can be used in finer-grained analyses to describe the temporal characteristics of phonatory versus non-phonatory (recovery) periods (Titze et al., 2007). Additional measures of vocal dose or vocal loading also can be derived from other monitored parameters, including cycle dose (total number of cycles of vocal fold vibration in a given period of time) and distance dose (estimate of the total distance traveled by the vibrating vocal folds in a given period of time based on a theoretical framework that combines values for f0, SPL, and phonation duration; Švec et al., 2003; Titze et al., 2003). The APM also provides simultaneous biofeedback capabilities (via a separate pager vibrator) based on thresholds set for f0 or SPL.

More recently, two additional ambulatory voice monitoring systems have been developed by other groups. The Vocalog Vocal Activity Monitor (Griffin Laboratories) is a commercially-available device that uses a neck-mounted contact microphone as the phonation sensor and provides simultaneous long-term monitoring and biofeedback for vocal SPL. The VoxLog (Sonvox) was recently introduced in Sweden as a research device (approval for clinical use is pending) that can provide long-term ambulatory monitoring of f0, SPL, and phonation duration using either of two sensor configurations: a neck-placed ACC for sensing the wearer's phonation, or a neck-mounted collar that contains both an ACC for sensing the wearer's phonation and a microphone for sensing background noise (environmental sound levels). A background noise measure is included in the VoxLog system under the assumption that this would be helpful in assessing the impact that such noise may have on the vocal SPL of the individual being monitored (e.g., the Lombard effect). The VoxLog also provides simultaneous biofeedback capabilities (via an integrated pager vibrator) based on thresholds set for f0 or SPL.

Future Directions

Current versions of ambulatory monitoring devices have begun to produce valuable new insights into daily patterns of voice use, and enthusiasm remains high for the potential of these devices to improve the diagnosis and treatment of voice disorders. The adoption of this technology into clinical practice, however, has been limited because of the lack of statistically robust studies to determine the true diagnostic capabilities of ambulatory voice measures. There is a need to collect long-term ambulatory data on groups of patients with voice-use-related disorders (before and after treatment) and matched controls that are large enough to permit robust statistical testing of derived measures. The ultimate goal is to determine if there are ambulatory voice measures that can reliably differentiate between pathological/hyperfunctional and normal patterns of daily voice use.

There is preliminary evidence that measures of vocal function in addition to f0, SPL, and phonation duration can be derived from the ACC signal. In particular, Zañartu (2010) has shown that a vocal system model can be used to estimate the glottal volume velocity airflow based on the acceleration signal from the anterior neck (skin) surface. Adding this capability to ambulatory voice monitoring could enhance its clinical use since it already has been shown that a combination of acoustic and aerodynamic (based on the glottal volume velocity waveform) measures have the potential to differentiate between hyperfunctional and normal patterns of vocal function.

Ambulatory biofeedback has been shown in early case studies to have some potential to facilitate vocal behavioral changes being targeted in voice therapy (Kay PENTAX, 2009). However, current devices are limited to rudimentary routines that only can provide feedback based on setting simple thresholds for f0 and SPL. Although apparently useful to some patients, this does not provide the capability to facilitate and reinforce changes in many other vocal behaviors that could be targeted in voice therapy. The potential for expanding the biofeedback capability of ambulatory voice monitoring systems could be increased dramatically once measures that differentiate pathological/hyperfunctional and normal patterns of vocal behavior are identified.

References

Airo, E., Olkinuora, P., & Sala, E. (2000). A method to measure speaking time and speech sound pressure level. *Folia Phoniatrica et Logopedia, 52*(6), 275-288.

Buekers, R., Bierens, E., Kingma, H., & Marres, E. H. (1995). Vocal load as measured by the voice accumulator. *Folia Phoniatrica et Logopedia, 47*(5), 252-261.

Cheyne. H. A. (2002). *Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck*. Unpublished doctoral dissertation, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA.

Cheyne, H. A., Hanson, H. M., Genereux, R. P., Stevens, K. N., & Hillman R. E. (2003). Development and testing of a portable vocal accumulator. *Journal of Speech, Language, and Hearing Research, 46*(6), 1457-67.

Hillman, R. E., Heaton, J. T., Masaki, A., Zeitels, S. M., & Cheyne, H. A. (2006). Ambulatory monitoring of disordered voices. *Annals of Otology, Rhinology, and Laryngology, 115*(11), 795-801.

Hillman, R. E., Holmberg E. B., Perkell J. S., Walsh M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech and Hearing Research, 32*(2), 373-392.

KayPENTAX. (2009). *Ambulatory phonation monitor: Applications for speech and voice.* Lincoln Park, NJ: Author.

Masuda, T., Ikeda, Y., Manako, H., & Komiyama, S. (1993). Analysis of vocal abuse: Fluctuations in phonation time and intensity in 4 groups of speakers. *Acta Otolaryngology, 113*(4), 547-552.

Nix, J., Svec, J. G., Laukkanen, A. M., & Titze, I. R. (2007). Protocol challenges for on-the-job voice dosimetry of teachers in the United States and Finland. *Journal of Voice, 21*(4), 385-396.

Ohlsson, A. C., Brink, O., & Lofqvist, A. (1989). A voice accumulation—validation and application. *Journal of Speech and Hearing Research, 32*(2), 451-457.

Popolo, P. S., Švec, J. G., & Titze, I. R. (2005). Adaptation of a pocket PC for use as a wearable voice dosimeter. *Journal of Speech, Language, and Hearing Research, 48*(4), 780-791.

Rantala, L., Vilkman, E., & Bloigu, R. (2002). Voice changes during work: Subjective complaints and objective measurements for female primary and secondary schoolteachers. *Journal of Voice, 16*(3), 344-355.

Ryu, S., Komiyama, S., Kannae, S., & Watanabe, H. (1983). A newly devised speech accumulator. *Journal for Oto-rhino-laryngology and its Related Specialties, 45*(2), 108-114.

Švec, J. G., Popolo, P. S., & Titze, I. R. (2003). Measurement of vocal doses in speech: Experimental procedure and signal processing. *Logopedics Phoniatrics Vocology, 28*(4), 181-192.

Švec, J. G., Titze, I. R., & Popolo, P. S. (2005). Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *Journal of the Acoustical Society of America, 117*(3), 1386-1394.

Titze, I. R., Hunter, E. J., & Švec, J. G. (2007). Voicing and silence periods in daily and weekly vocalizations of teachers. *Journal of Acoustic Society of America, 121*(1), 469-478.

Titze, I. R., Švec, J. G., & Popolo, P. S. (2003). Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues. *Journal of Speech, Language, and Hearing Research, 46*(4), 919-932.

Verdolini, K., & Ramig, L. O. (2001). Review: Occupational risks for voice problems. *Logopedics Phoniatrics Vocology, 26*(1), 37-46.

Verdolini, K., Rosen, C., & Branski, R. C. (2005). *Classification manual for voice disorders-I, Special Interest Division 3, Voice and Voice disorders, American Speech-Language Hearing Division.* Mahwah, NJ: Lawrence Erlbaum.

Watanabe, S. T., Oda, M., Fukaura, J., & Komiyama, S. (1987). Measurement of total actual speaking time in a patient with spastic dysphonia. *Folia Phoniatrica, 39*(2), 65-70.

Watanabe, K. S., Ryu, S., & Kannae, S. (1984). A newly designed automatic phonometer. *Agressologie, 25*(9), 1035-1037.

Zañartu., M. (2010). *Acoustic coupling in phonation and its effect on inverse filtering of oral airflow and neck surface acceleration.* Unpublished doctoral dissertation, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.

Zañartu, M., Ho, J. C., Kraman, S. S., Pasterkamp, H., Huber, J. E., & Wodicka, G. R. (2009). Air-borne and tissue-borne sensitivities of bioacoustic sensors used on the skin surface. *IEEE Transactions on Biomedical Engineering, 56*(2), 443-551.